Real-time motion detection using an analog VLSI zero-crossing chip

Wyeth Bair, Christof Koch

Wyeth Bair, Christof Koch, "Real-time motion detection using an analog VLSI zero-crossing chip," Proc. SPIE 1473, Visual Information Processing: From Neurons to Chips, (9 July 1991); doi: 10.1117/12.45541

SPIE.

Event: Orlando '91, 1991, Orlando, FL, United States

Real-time motion detection using an analog VLSI zero-crossing chip

Wyeth Bair[1,2] and Christof Koch[1]

[1]California Institute of Technology, Computation and Neural Systems Program, 216-76,
Pasadena, California 91125
[2]Hughes Aircraft Artificial Intelligence Center,
Malibu, California 90265

ABSTRACT

We have designed and tested a one-dimensional 64 pixel, analog CMOS VLSI chip which localizes intensity edges in real-time. This device exploits on-chip photoreceptors and the natural filtering properties of resistive networks to implement a scheme similar to and motivated by the Difference of Gaussians (DOG) operator proposed by Marr and Hildreth (1980). Our chip computes the zero-crossings associated with the difference of two exponential weighting functions and reports only those zero-crossings at which the derivative is above an adjustable threshold. A real-time motion detection system based on the zero-crossing chip and a conventional microprocessor provides linear velocity output over two orders of magnitude of light intensity and target velocity.

1. INTRODUCTION

The zero-crossings of the Laplacian of the Gaussian, $\nabla^2 G$, are often used for detecting edges. Marr and Hildreth (1980) argued that the Mexican-hat shape of the $\nabla^2 G$ operator can be approximated by the difference of two Gaussians (DOG). In this spirit, we have built a chip that takes the difference of two resistive-network smoothings of photoreceptor input and finds the resulting zero-crossings. The Green's function of the resistive network, a symmetrical decaying exponential, differs from the Gaussian filter. Figure 1 shows the "Mexican-hat" shape of the DOG superimposed on the "witch-hat" shape of the difference of exponentials (DOE) filter implemented by our chip.

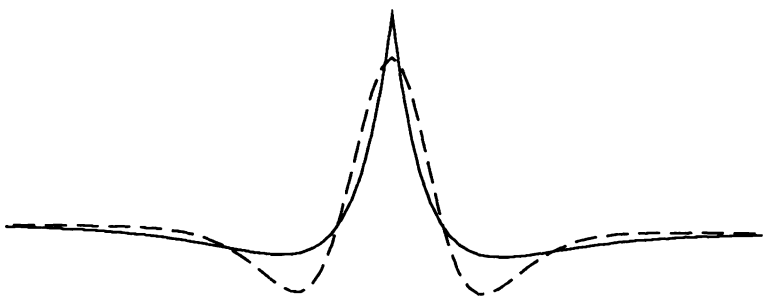

Fig. 1. The Mexican-hat shape of the difference of Gaussians (dotted) and the witch-hat shape of the difference of exponentials (DOE) filter implemented by our chip.

This implementation has the particular advantage of exploiting the smoothing operation performed by a linear resistive network, shown in Figure 2. In such a network, data voltages d are applied to the nodes along the network via conductances G, and the nodes are connected by resistances R. Following Kirchhoff's laws, the network node voltages v settle to values such that power dissipation is minimized. One may think of the network node voltages v as the convolution of the input with the symmetrical decaying exponential filter function. The characteristic length of this filter function is approximately $1/\sqrt{RG}$, where G is the data conductance and R the network resistance.

Such a network is easily implemented in silicon and avoids the burden of additional circuitry which others have used to implement Gaussian kernels. Our simulations with digitized camera images show only minor differences

0-8194-0582-5/91/$4.00

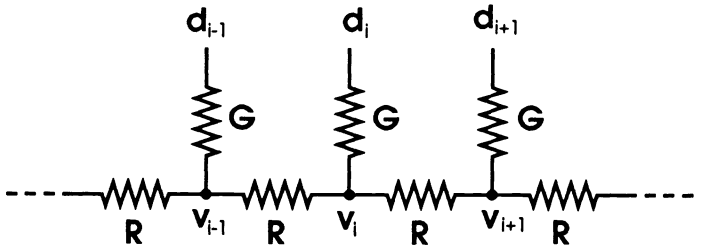

Fig. 2. In this 1-D resistive network, data voltages d are applied to the network via conductances G. Resistances R couple neighboring nodes in the network.

between the zero-crossings from the DOE filter and those from the DOG.

2. ANALOG VLSI IMPLEMENTATION

This chip was implemented with a 2.0μm CMOS n-well process available through the MOSIS silicon foundry. Intensity edges are detected using four stages of circuitry: photoreceptors capture incoming light, a pair of 1-D resistive networks smooth the input image, transconductance amplifiers subtract the smoothed images, and digital circuitry detects zero-crossings. Figures 3 and 4 show block diagrams for two pixels of the 64 pixel chip.

Processing begins at a line of photoreceptors spaced 100μm apart which encode the logarithm of light intensity as a voltage VP, shown in Figure 3. The set of voltages from the photoreceptors are reported to corresponding nodes of two resistive networks via transconductance amplifiers connected as followers. The followers' voltage biases, V_{G1} and V_{G2}, can be adjusted off-chip to independently set the data conductances for each resistive network. The network resistors are implemented as Mead's saturating resistors (Mead, 1989). Voltage biases V_{R1} and V_{R2} allow independent off-chip adjustment of the two network resistances. The data conductance and network resistance values determine the space constant of the smoothing filter which each network implements. The sets of voltages $V1$ and $V2$, shown in Figure 3, represent the two filtered versions of the image. Wide-range transconductance amplifiers (Mead, 1989) produce currents, I, proportional to the difference $V1 - V2$.

Figure 4 shows the final stage of processing which detects zero-crossings in the sequence of currents I and implements a threshold on the slope of those zero-crossings. Currents I_i and I_{i+1} charge or discharge the inputs of an exclusive OR gate. The output of this gate is the first input to a NAND gate which is used to implement the threshold. A current proportional to the magnitude of the difference $I_i - I_{i+1}$ charges the second input of the NAND gate, while a threshold current discharges this input. If the charging current, representing the slope of the zero-crossing, is greater than the threshold current set off-chip by the bias voltage V_{thresh}, this NAND input is charged to logical 1, otherwise, this input is discharged to logical 0. The output of the NAND gate, VZ_i indicates the presence, logical 0, or the absence, logical 1, of a zero-crossing with slope greater than I_{thresh}.

A final stage of circuitry is used to multiplex the sequence of 63 bits, VZ, and corresponding currents $I_i - I_{i+1}$ indicating the slope of the zero-crossings.

3. BEHAVIOR

We tested the behavior of the chip by placing a small lens above the silicon wafer to focus an image onto the array of photoreceptors. The input light profile that we used is shown in Figure 5a. Figure 5b is an oscilloscope trace showing the smoothed voltages ($V1$ and $V2$ of Figure 3) corresponding to the filtered versions of the image. The difference of these two smoothed voltage traces is shown in Figure 5c. Arrows indicate the locations of two zero-crossings which the chip reports at the output. The reported zero-crossings accurately localize the positions of the edges in the image. The trace in Figure 5c crosses zero at other locations, but zero-crossings with slope less than

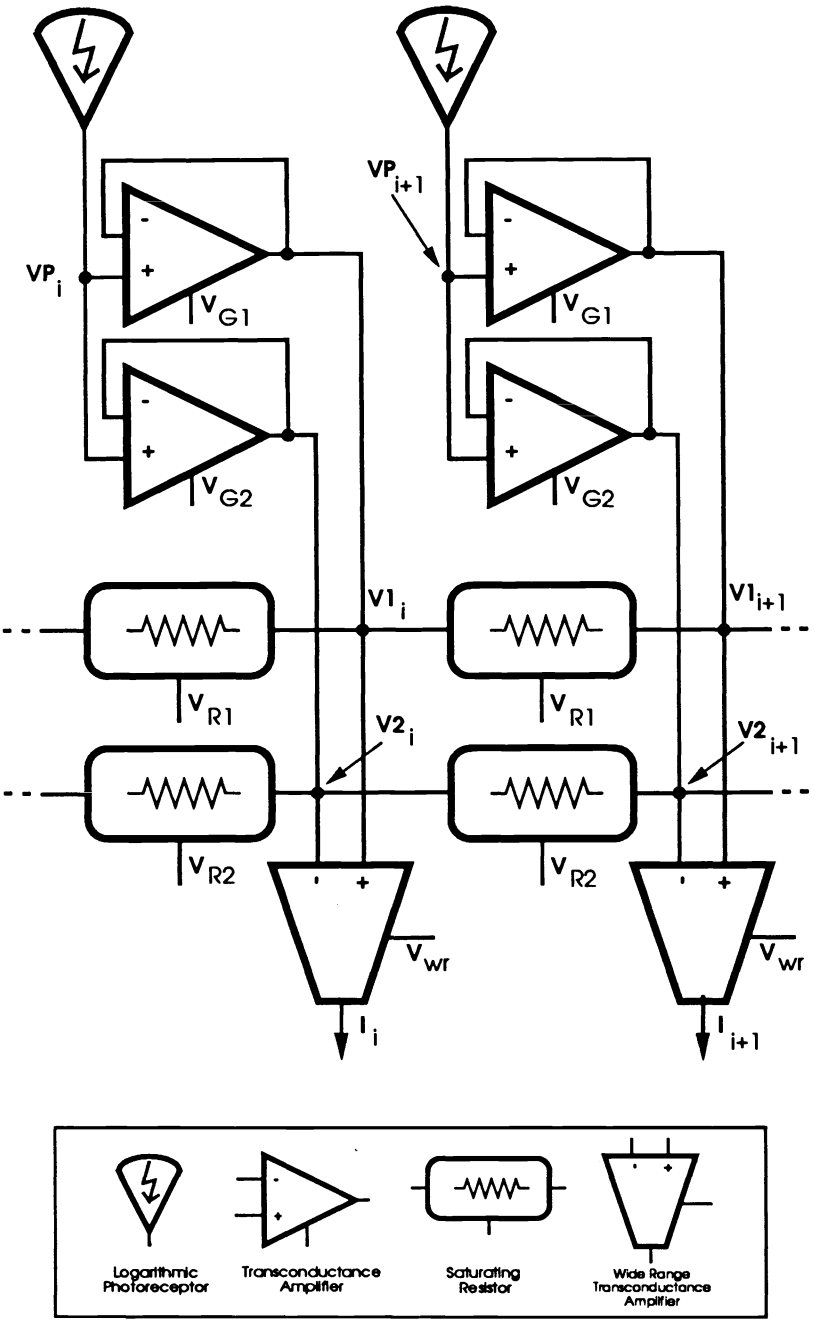

Fig. 3. Zero-crossing chip circuit diagram. Logarithmic photoreceptors encode light intensity as voltages, VP, which are reported to the nodes of two resistive networks via transconductance amplifiers connected as followers. The voltage biases VG set the conductances. The network resistances, $R1$ and $R2$, are implemented as saturating resistors and are also adjustable from voltage biases. The filtered images are subtracted by wide-range transconductance amplifiers which ouput currents, I, proportional to the voltage across their inputs.

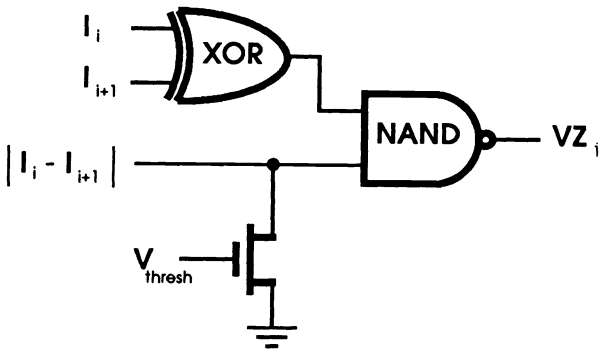

Fig. 4. An exclusive-OR gate is used to detect a zero-crossing, and a transistor shunts current to threshold on the magnitude of the derivative. The conjunction of a zero-crossing and a large derivative cause an edge to be reported at the final output.

the adjustable threshold are masked by the circuitry shown in Figure 4. This allows for noise and imperfections in the circuitry and can be used to filter out weaker edges which are not relevant to the application.

Figure 6 shows the response when two fingers are held one meter from the lens and swept across the field of view. The fingers appear as bright regions against a darker background. The chip accurately localizes the four edges (two per finger) as indicated by the pulses below each voltage trace. As the fingers move quickly back and forth across the field of view, the image and the zero-crossings follow the object with no perceived delay. From sequences of frames like these, we can compute optical flow. Note that these are not successive frames, but are more representative of every hundredth frame that the motion detection system will receive.

4. MOTION FROM ZERO-CROSSINGS

The motion detection system consists of one zero-crossing chip interfaced to a 12.5MHz 80286 microprocessor-based single-board computer. The interface allows the microprocessor to receive 63-bit frames of zero-crossing data at just over 320 frames per second. As each new frame is read, the microprocessor updates the cumulative displacement of each zero-crossing and increments the number of frames over which that displacement has occurred. The system assumes that zero-crossings will not move more than 2 pixels per frame. With our optics, this assumption is violated only at velocities in excess of approximately 700 degrees per second.

After tracking zero-crossings for a fixed number of frames, their individual velocities are computed in pixels per frame. These velocities are averaged for all zero-crossings which have been tracked for longer than a fixed number of frames. For the data shown here, an average full-field velocity is reported every second. Figure 7 shows the average and standard deviation of the reported velocity over a one minute period for input velocities ranging in magnitude from zero to 450 pixels per second at two light levels. The dotted lines show the standard deviation of the output velocity. Over most of this range, the standard deviation was less than four percent of the average value. Image velocity was limited by the lens and stimulus. The data shown for 10 W/m^2 is representative of the system response for light levels of 1 W/m^2 and higher. Below 1 W/m^2 the zero-crossing chip was unable to localize higher velocity edges. We believe this is due to R-C time constants associated with the circuitry of the analog chip. Also, as seen in Figure 7, the reported velocity is less than the image velocity but remains linear. At lower light levels, zero-crossings due to offsets are more prevalent and introduce zeros into the average velocity computation, thus lowering the reported velocity. Such spurious zero-crossings can undermine the accuracy of the average velocity in more subtle ways as well. As light intensity drops, the linear range of output for this system becomes smaller around zero. Below 100 mW/m^2, the zero-crossing chip fails to detect edges, and the system cannot even detect

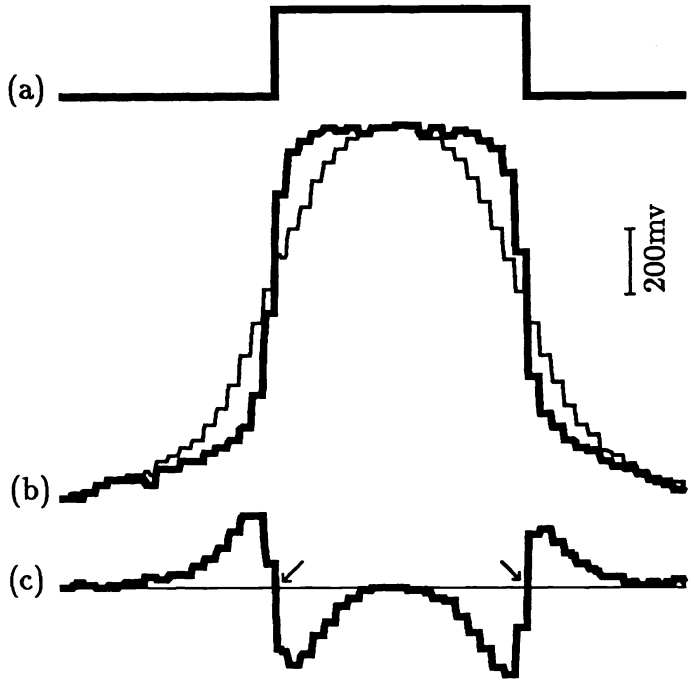

Fig. 5. Response of the zero-crossing chip to a light bar stimulus. (a) Input light intensity, (b) voltage traces from the two resistive networks, and (c) difference of voltage traces and arrows indicating the locations of zero-crossings localized by the chip. The threshold suppresses zero-crossings having derivatives of small magnitude.

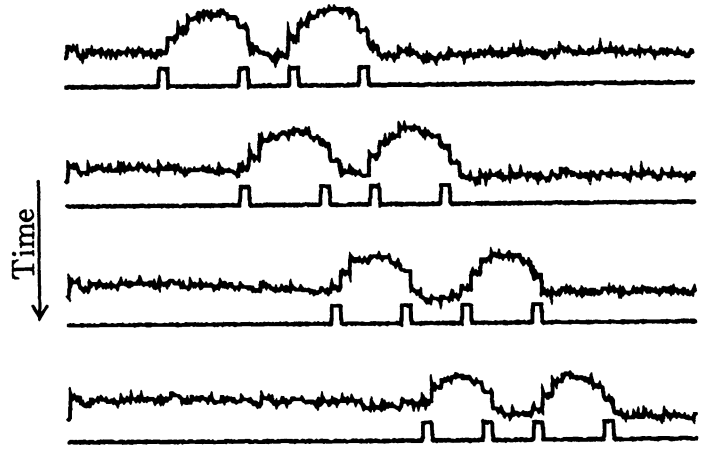

Fig. 6. Zero-crossing chip response as two fingers are waved one meter in front of the lens. The upper traces show voltages from one resistive network; the lower traces show positions of zero-crossings reported by the chip.

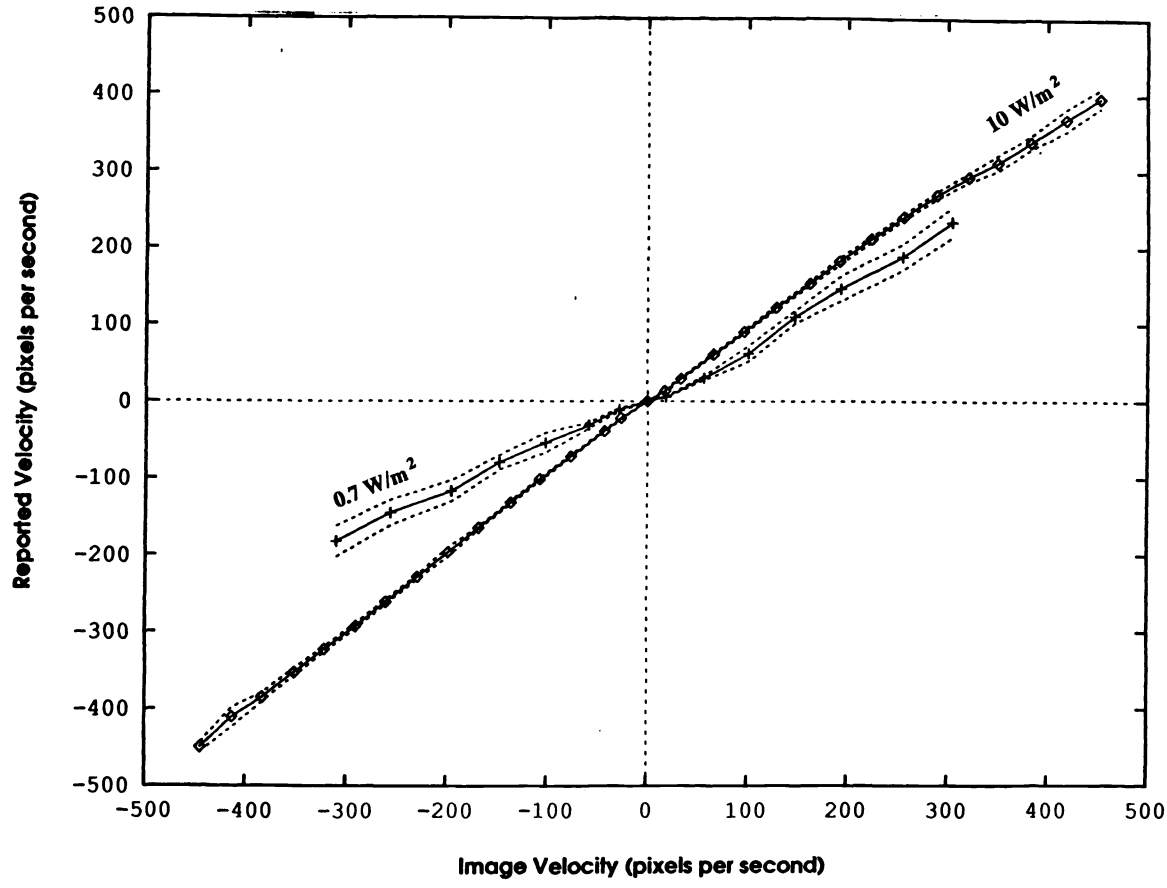

Fig. 7. Zero-crossing motion detection system output for two light intensities. At light intensities above 1 W/m^2, the output is linear and accurate over a large range of velocities. At lower intensities, the zero-crossing chip cannot localize fast edges, and lower signal-to-offset ratios introduce spurious zero-crossings that compromise accuracy. (Dotted lines show standard deviation.)

direction of motion. Qualitatively, the useful range of operation for this system is from bright sunlight to dim indoor fluorescent or incandescent lighting, and this is achieved without changing parameters.

The zero-crossing chip fails at low light and contrast levels due to the small signal-to-offset ratio. Imperfections in the fabrication process cause many of the signals in the analog chip to be corrupted. The magnitude of this noise, called offsets, is a substantial fraction of the magnitude of the signal reported by the logarithmic photoreceptors. Although the logarithmic receptor allows operation over a wide range of lighting conditions, it compresses the range of voltages which are used to encode any particular scene and therefore decreases the signal-to-noise ratio. A hysteretic photoreceptor similar to the one used in the second chip described in this paper would improve the signal-to-noise ratio, but would also increase sensitivity to lighting changes, and possibly compromise sensitivity to small velocities.

Another limitation on the performance of the zero-crossing chip is the photoreceptor response time. The measured response time of the chip to the appearance of a detectable discontinuity in light intensity varies from about 100μsec in bright indoor illumination to about 10msec in a dark room, and these response times seem to be dominated by the logarithmic photoreceptor.

Finally, spatial and temporal aliasing may limit the performance of this system. As the spatial frequency of

features increases, zero-crossings appear closer together and the correspondence problem arises. This is a function of the environment, the lens, and the photoreceptor spacing on the chip. Interfacing the zero-crossing chip to a digital computer requires clocking the output from the chip. In theory, this causes temporal aliasing at higher velocities, but the slow time response of the photoreceptors cause the system to fail before temporal aliasing is noticed.

5. CONCLUSION

Our analog VLSI chip demonstrates that finding the thresholded zero-crossings of the difference of exponential filters is a robust technique for localizing intensity edges in real-time. This supports the approach of compromising optimality of an algorithm for compactness and simplicity of implementation. The motion detection system based on thresholded zero-crossings produced linear output over two orders of magnitude of light intensity and target velocity. Again, this shows the usefulness of implementing simple algorithms in analog VLSI and encourages us to continue producing devices which encroach on the computational domain of larger general purpose digital processors. We are currently integrating all of the processing for motion detection onto single analog chips.

6. ACKNOWLEDGMENTS

We thank Carver Mead for providing laboratory resources for the design, fabrication, and initial testing of this chip. Thanks also to Steve DeWeerth, Misha Mahowald, and John Harris for their help throughout the years. Our laboratory is partially supported by grants from the Office of Naval Research, the Rockwell International Science Center and the Hughes Aircraft Artificial Intelligence Center. Wyeth Bair is supported by a National Science Foundation Graduate Fellowship and performed some of this work at the Hughes Aircraft AI Center.

7. REFERENCES

1. D. Marr and E. C. Hildreth, "Theory of edge detection," *Proc. Roy. Soc. Lond. B* **207:** 187–217, 1980.
2. C. A. Mead, *Analog VLSI and Neural Systems,* Addison-Wesley, Reading, MA, 1989.